DOCUMENT RESUME

ED 157 897 TM 006 513

AUTHOR Denton, Jon J.; Crowley, Lee B.
TITLE The Relation Between Unit Test and Retention Test Performances: A Case for the Personalized System of Instruction Approach to Grading.
PUB DATE [76]
NOTE 15p.; Page 14 may be marginally legible

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Academic Achievement; Achievement Tests; Correlation; Course Objectives; Criterion Referenced Tests; Grading; Higher Education; *Individualized Instruction; *Mastery Learning; *Mastery Tests; *Retention; *Testing; *Units of Study
IDENTIFIERS *Final Examinations

ABSTRACT

To determine the correlation between unit test performance and retention on final examinations, scores of 68 secondary level teaching candidates enrolled in a general methods course were studied. Unit tests were administered frequently to determine student achievement of the 38 course objectives. Two retests per student were allowed, with retesting focusing only on those objectives not achieved. The 83-item final examination was categorized with respect to the 10 instructional units in the course. Results indicated a correlation of .406 between unit test performance (expressed as objectives achieved) and retention scores. Correlation coefficients between unit objectives achieved (both total number and number achieved on the initial attempt) and the retention test subscore for that unit were also positive. In conclusion, empirical evidence supported the assignment of grade credit on the basis of attainment of objectives during instruction; however, the final examination should also contribute to the course grade. Such grading policy favors the Personalized System of Instruction. The correlation values are not so great as to encourage the Mastery Learning approach which suggests that grades can be assigned entirely by final exam performance. (CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED157897

The Relation Between Unit Test and Retention Test Performances: A Case for the Personalized System of Instruction Approach to Grading.

Jon J. Denton, Lee B. Crowley

Texas A&M University

The purpose of this investigation was to determine whether accomplishment of performance objectives determined by satisfactory performance on unit tests during instruction, is related to the candidate's performance on an end-of-course retention test. Investigating this relation provided empirical evidence supporting the assignment of grade credit on the basis of objective attainment during instruction rather than resorting to the final test to determine a student's course grade. This investigation was conducted with sixty-eight candidates in professional education coursework, which immediately preceded student teaching.

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Jon Denton

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

TM006 513

The Mastery Learning strategy[1] and the Personalized System of Instruction[2] have provided the rationale for numerous investigations on individualized instruction over the past decade. The following paragraphs provide brief descriptions of these approaches which presume to meet the instructional needs of each student.

The Mastery Learning strategy limits mastery of the subject to the behavior component of stated performance objectives and provides each learner with the necessary time and instruction to achieve those objectives. Instructional units are approximately two weeks in length and sequential, that is, early units provide prerequisite material for later units. The instructional mode endorsed by this approach includes large group instruction with the instructional activities and time management under the teacher's control. Criterion referenced unit tests are administered at the conclusion of each unit to measure student progress and success in achieving the performance objectives. Ordinarily, this test is corrected by students but not recorded for a grade, the results are keyed to the objectives and used by the student as a guide to review for the final test. For those who do not reach mastery, several remediation techniques are recommended. The use of a variety of remediation materials and activities is felt to meet the instructional needs of the students more adequately than those he experienced initially in the unit. The course continues through succeeding units with assessments and prescriptive remediations until each student has mastered all the objectives in the course. A final test (summative examination) is administered at the conclusion of the course for grade assignment purposes.[3]

In comparison to the mastery plan, a course based on the Personalized System of Instruction is typically divided into several units each of which

centers around a small number of performance objectives to be attained in one week.. The instructor prepares study guides which contain questions and comments related to the textbook materials and performance objectives. Lectures and demonstrations, if used with this approach, serve to motivate and create interest rather than provide substantive information. Typically, a student reads the unit materials and responds to the study-guide questions. He then comes to class, where he completes a written unit test referenced to the unit objectives. A proctor scores the examination and provides immediate feedback. If the student has answered enough questions to reach criterion, usually above 80%, he can proceed to the next unit at his own pace. However, if he fails to reach the criterion level stated in the objectives, he restudies the original materials and retakes alternate forms of the unit test until he does reach criterion. Grades are based on the number of units (often expressed in terms of objectives) mastered with a small weight on final examinations[4].

Reviews of the research on these individualized approaches to instruction have addressed cognitive outcomes (achievement, retention, transfer) and affective outcomes (interest, attitudes) of learners[5, 6, 7]. These reviews conclude Mastery Learning and the Personalized System of Instruction approaches enhance cognitive outcomes and cultivate positive attitudes. Moreover, these reviews have examined special components of each system, for example, pacing of instruction, use of proctors, use of lectures, criterion levels for mastery, testing format, frequency of testing, grading practices, and interactions between the instructional component and student learning style to determine the respective influence of each component on the positive findings reported in the literature.

Although these reviews have examined the frequency of testing and the nature of examinations, the relation between unit test performances and retention test results has not been addressed. A significant relation between these assessment variables would give credence to the practice of awarding grade credit on the basis of mastery performances on unit tests. Given the emphasis on instructional remediation by both of these approaches to individualized instruction, a case can be made that grade credit should be awarded whether the mastery performance occurred on the first or successive administrations of the unit test. With this antecedent, the research question for this investigation became: Do positive relations exist between unit test results, whether from initial or retest administrations, and retention test performances (total score, and unit subscores)?

Methodology

Sample

This investigation was conducted under the sponsorship of an educational curriculum and instruction department at a Land Grant University accredited by the Southern Association of Colleges and Universities and the National Council for Accredition of Teacher Education.

Sixty-eight secondary level teaching candidates enrolled in a general teaching methods course participated in this investigation. Academically, each candidate met the following departmental criterion for admission into student teaching to be enrolled in the class:

a) Overall Grade Point equal to or exceeding 2.25 based on a 4 point scale.

b) Grade Point Ratio from 18 semester hours of teaching field coursework equal to or exceeding 2.25 based on a 4 point scale.

c) Senior standing in the University, i.e., completed 95 semester hours.

d) Endorsement by a professor in teaching field attesting to the candidate's competence in that field.

Certainly no claim is made about random selection of candidates, since all students who completed the course and retention test were included in the sample. This decision was made to maximize the sample realizing that reported alpha levels, given this assignment procedure, increase the likelihood of type II errors.

Course Description

The independent variables under consideration in this investigation were obtained as part of the instructional process for the general methods course, Principles and Practices of Teaching (EDCI 423). This course consisted of ten instructional units which were developed to encourage independent study and self-paced learning. Thirty-eight performance objectives delineated the expected cognitive skills and the levels of performance (usually 80%) necessary to signify their acquisition for each of the various units.

The instructor was available during the scheduled course hours and during his posted office hours for individual tutorial sessions and small group discussions. However, with the exception of three units, class attendance was not required. The units which required attendance involved a number of group activities and demonstrations involving

communication skills.

Unit tests and assessments for objectives were administered in accordance with the chronological schedule accompanying the course syllabus. Students were given immediate feedback on the results of examinations and encouraged to review both the material they initially experienced and supplementary material if they failed to reach criterion on one or more of the unit objectives. A limit of two retests per student was possible for each unit, with retesting focusing on only those objectives not achieved. Monitoring of an individual's progress was maintained by utilizing a computer support system which scored, recorded and provided printouts reflecting whether objectives were achieved on the first or succeeding trials[8].

Course grades were determined on the basis of the number of objectives achieved with no penalty being assessed because an objective was not accomplished on the initial attempt. In order to qualify for an "A" grade, each student had to submit assignments and attempt all thirty-eight objectives, attaining ninety percent (34) of the objectives, the minimum number for an "A" grade, would not suffice. A final examination (retention test) was administered at the conclusion of the course to determine the strengths and weaknesses of the instructional program. Students were instructed not to prepare for this examination, since it had no bearing whatsoever on their course grades.

Instrumentation

This final examination contained eighty-five items of three types, i.e., multiple choice, true-false and classification. Items were categorized with respect to the ten instructional units with three units

(Discipline, Performance Objectives, Test Construction) being represented by fifteen items each, four other units (Interaction Analysis, Teacher Questions, Grading Practices, Teacher/Law) were each represented by ten items and three of the ten units (Unit Planning, Simulations/Games, Professional Materials) were not represented on the examination. A KR-20 estimate of internal consistency for the examination was determined to be .840.

Content validity of the instrument was addressed by selecting items from unit tests which were referenced to the performance objectives of the course. The first draft contained one hundred twenty-five items which represented nearly every item contained in the numerous unit tests. Pilot testing of this instrument revealed a number of items with little discriminating ability and others which contained ambiguous stems and answer choices. After review and appraisal, forty items were removed resulting in the final form containing eighty-five items.

Unit tests were administered frequently to determine student progress and accomplishment of the performance objectives for the course. These tests usually provided five items/objective and required four correct responses to signify objective attainment. Three versions of each unit test were prepared since that number of testing opportunities were possible. However, few instances were recorded where three tests were taken by a student with respect to a particular objective. Item statistics and estimates of internal consistency were not determined for the unit examinations, but content validity was considered by both instructor (before testing) and student (during testing) alike since each assessed the relation between test items and the stated objective. If students challenged

the relation of a particular item to an objective, then efforts were made to explain the relation. The outcome of the challenge resulted either in the retention or removal of the item depending on the situation.

Findings

Unit test performances of students expressed in terms of unit objectives achieved whether on the first, second, or third attempt were obtained from printouts of the computer support system designed for this course. Values gleaned from this support system indicate that on the average, each candidate achieved 33.8 objectives while remediating five objectives.

Retention test data were obtained from the administration of the retention test at the conclusion of the course. The total score and unit subscores were expressed both as raw scores and standardized T scores. Standardized scores were selected for analysis in the investigation, consequently the means and standard deviations for each set of retention scores were identical, 50 and 10, respectively.

A scatter plot was generated and the corresponding correlation coefficient was determined between the total number of objectives achieved by each student and his performance (total score) on the retention test using statistical routines from SAS[9]. Figure 1 presents the scatter plot which graphically depicts the correlation of .406 (p .0006) between unit test performance expressed as objectives achieved and retention scores as structured by this investigation. The scatter plot illustrates

Insert Figure 1

that a high proportion of students achieved thirty-four or more objectives. A factor which undoubtedly influenced the distribution of objectives achieved was the grading system which required the attainment of thirty-four objectives for an "A" grade. Other cogent observations from the plot are: the greatest range of retention scores (30.7 to 69.3) occurred when thirty-four objectives were achieved; and higher retention scores occurred as more objectives were achieved.

Additional analyses were conducted to determine if the relation observed between the cumulative values of objectives achieved and the retention test would be observed when individual instructional units were examined. Units were selected for additional analyses if they were developed around five or more performance objectives. Four of the units met this criterion (Discipline, Performance Objectives, Test Construction, Grading Practices) and were subsequently analyzed. Correlation coefficients were determined between the unit objectives achieved, both the total number and the number achieved on the initial attempt, and the retention test subscore for that unit. The results of the analyses are presented in Table 1.

Insert Table 1

Clearly, the relation determined from the overall calculation is reflected in three of the four units. One possible explanation for the smaller coefficients for the performance objectives unit is that nearly all of the test items related to this unit required the student to classify behavioral phrases of objectives into the various categories of the cognitive and affective domains. The instructional activities developed to impart these skills were inadequate and subsequently were

removed as were the performance objectives they related to. Unfortunately, the objectives and activities were included in the instructional program for this investigation, however.

Examining the correlations more carefully, it is noted the correlation values for the "total objectives achieved" and the coefficients for the "objectives achieved on the first attempt" are nearly equivalent. Apparently the retention of cognitive skills is not too different for those students who achieved the objectives on the first trial compared to those who remediated and retested.

Discussion

These findings permit an affirmative response to the research question under consideration in this investigation. Clearly, a positive relation was found to exist between unit test performance and retention scores. Moreover, the bivariant distribution depicts greater variation among retention test performances as the number of objectives achieved (unit test performances) increase. This variation of scores might suggest that individuals who required more than one attempt to reach criterion on unit objectives would not perform as well on the retention test as those who achieved the objectives on an initial attempt. If this explanation is viable, then higher correlation values should occur between unit retention scores and objectives achieved on the first attempt then when total objectives achieved is one of the variables. This phenomenon did not occur consistently, thus the explanation is not acceptable. Apparently students given a "second chance" on a unit test tended to perform as well on the retention test as those who achieved the objectives without remediation.

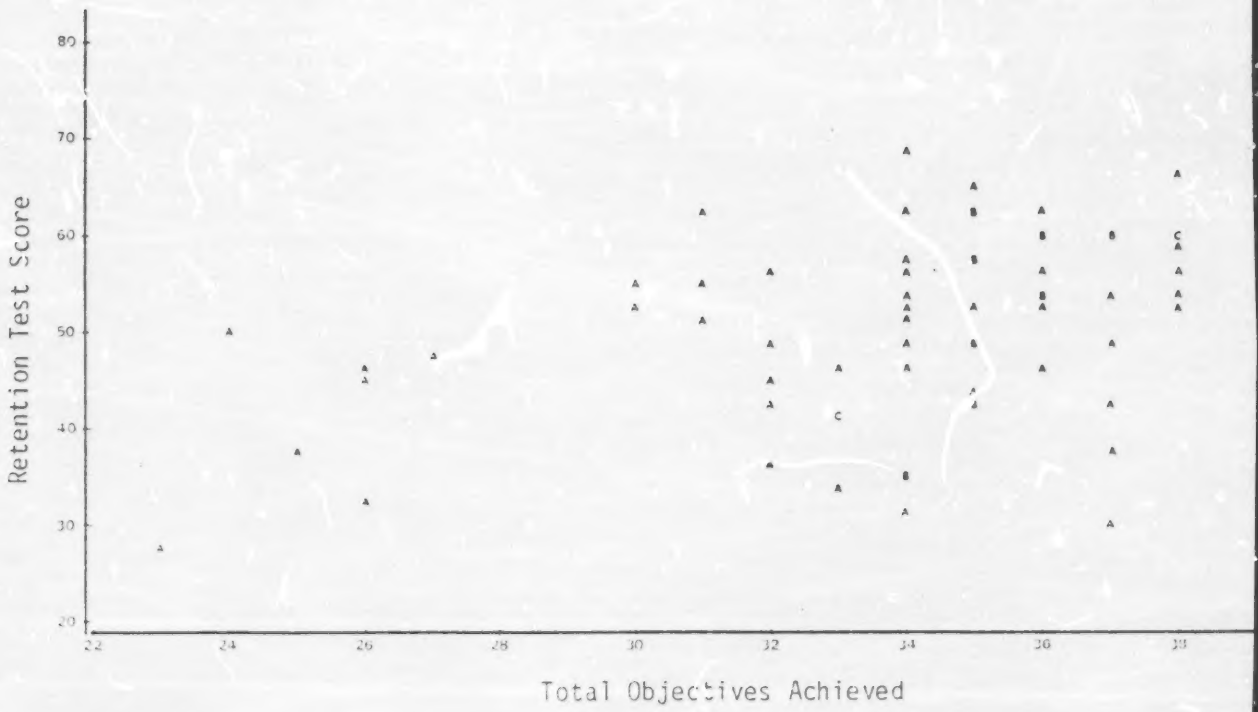
The primary goal for undertaking this investigation was to determine if empirical support is possible for assigning grade credit on unit test performances, given an individualized instructional approach which includes remediation and retesting of unit objectives not achieved on an initial test. Since positive correlations were found between objectives achieved and retention test performances for both cumulative values and unit values, empirical evidence does exist to support this goal. Certainly the correlation values are not so great as to encourage adoption of a grading policy based solely on objective attainment (unit test performance). However, these results do support the notion of assigning grade credit for objective attainment with the final test of the course contributing to the course grade. This recommendation is consonant with the grading system of the Personalized System of Instruction[10]. Under this system, student reinforcement is provided through knowledge-of-results and grade credit being awarded for successful performances on the unit examinations. Conversely, the Mastery Learning approach suggests that grade assignment be made entirely on the basis of final test performance[11]. Reinforcement under this system appears to be predicated entirely on knowledge-of-performance.

Pragmatically, student motivation appears to be favored by the grading approach endorsed by the Personalized System of Instruction. Moreover, this investigation has provided empirical evidence which are consistent with this grading policy. Given these observations, individualized systems of instruction which include remediation and retesting in their instructional design are encouraged to assign grade credit on successful unit test performance.

Table 1

Correlation Coefficients Between Standardized Retention Scores / Unit and Objectives Achieved

		Objectives Achieved all attempts	Objectives Achieved first attempt
Discipline Unit	r	.231	.258
Retention Score		.058	.027
Performance Objectives	r	.175	.084
Unit Retention Score		.152	.496
Test Construction	r	.355	.303
Unit Retention Score		.003	.012
Grading Practices	r	.307	.466
Unit Retention Score		.011	.0001



legend A=1 observation
B=2 observations
C=3 observations

Figure 1
Bivariant Distribution of Student Performance

References

1. Benjamin S. Bloom, "Learning for Mastery," Evaluation Comment, 1 (1968).

2. Fred S. Keller, "Goodbye, Teacher . . ." Journal of Applied Behavior Analysis, 1 (1968) pp. 79-89.

3. James H. Block (Ed), Schools, Society and Mastery Learning (New York:Holt, Rinehart & Winston, 1971) pp. 16-18.

4. Arthur L. Robin, "Behavioral Instruction in the College Classroom," Review of Educational Research 46(1976) pp. 314-317.

5. James H. Block (Ed), Mastery Learning: Theory and Practice, (New York:Holt, Rinehart & Winston, 1971).

6. Schools, Society and Mastery Learning, pp. 28-64.

7. "Behavioral Instruction in the College Classroom," pp. 320-343.

8. Jon J. Denton, Roy M. Bennett, "A Computer Support System for a CBTE Program," Educational Technology, 15 (1975) pp. 50-53.

9. Arthur J. Barr, H. Goodnight, John P. Sall, J. T. Helwig, A User's Guide to SAS 76, (Raleigh, N.C.:SAS Institute Inc., 1976).

10. "Behavioral Instruction in the College Classroom," pp. 314-315.

11. Benjamin S. Bloom, J. Thomas Hastings, George F. Madaus, Handbook on Formative and Summative Evaluation of Student Learning (New York: McGraw-Hill Book Co., 1971) pp. 53-56.